

計画系論文における研究キーワードの構造化に関する研究

Structuring of Research Key Word in Proceedings of Planning Study

中岡良司*・今尚之**・佐藤 肇***

by Ryoji NAKAOKA, Naoyuki KON, Keiichi SATOH

1. はじめに

本研究は、1970（昭和45）年から1991（平成3）年までの土木学会年次学術講演会（以下、年講）第IV部門の発表論文 4,835件を対象に、論文表題から研究キーワードを自動抽出し、その構造化を図るとともに年代別の関心を明らかにしたものである。

年講第IV部門はいわゆる計画系論文の発表部門であり、自由投稿であることや頁数も2頁と少ないことから、全国の研究者の最新研究あるいは萌芽的研究を網羅している。そこで、本研究では年講論文を対象に、①計画系論文の主要な関心を論文の表題から明らかにすること、②使用されている研究キーワードの関連を構造的に把握すること、③過去約20年間の研究キーワードの変遷を分析することを研究の目的とした。

2. データベース構築と研究フロー

データベースは、その内容は当然のこととして、量的に満足できるものでなくては価値がない。今日、データベース構築の大きな障害は、コンピュータへのデータ入力作業である。そこで、本研究で対象とした年講第IV部門の論文 4,835件の入力に際しては、パソコンによるOCRシステムを利用し、約2週間で入力を完了した。OCRとは Optical Character Recognition（光学式文字認識）の略であり、印刷物あるいは手書き原稿から文字を認識してコンピュータで扱える文字コードに変換するシステムである。本研究では、解像度320dpiのスキャナおよび市販ソフト「ワードアイ」（マクロエス(株)）を使用した。

キーワード：論文表題・キーワード分析

* 正会員 北見工業大学助手 土木開発工学科
(〒090 北見市公園町165)

** 正会員 工博 小樽商科大学助手

*** 正会員 工博 北海道大学教授

本システムの読み取り精度は約91%であった。詳細は文献1)を参照されたい。

構築したデータベースにより、発表論文件数の推移を図-1に示す。1970年（第25回大会）には107件であった論文数は1991年（第46回大会）には349件となり約2.4倍に増加し今後も増加の傾向にある。

研究フローを図-2に示す。構築したデータベースに対し、第1段階では、かな漢字分離方式によるプログラムでキーワードの自動抽出を行った。キー

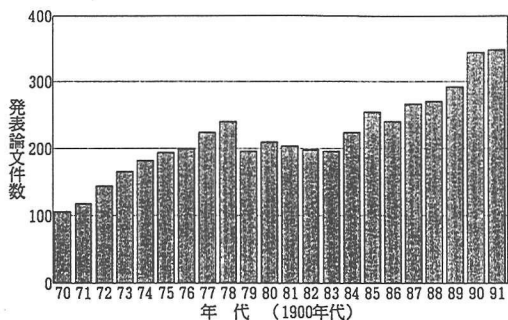


図-1 年講（第IV）発表論文の推移

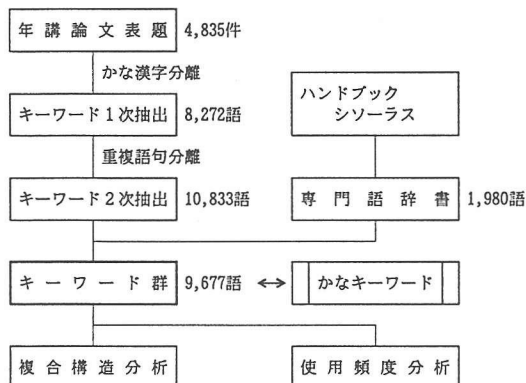


図-2 研究フロー

ワードとはその論文を特徴づける主要な用語のことである。かな漢字分離方式とは、かなと漢字が混在している論文表題から2文字以上の漢字列を取り出す方法であり、キーワードの多くが2文字以上の漢字から構成されている点に着目し独自に考案したものである。ただし、「乗り継ぎ」などこの原則には当てはまらない用語もあるため後に補完する対策を取った。第2段階では、前段階で得られたキーワード群に対し、相互に重複語句を検査するプログラムで新たなキーワードの抽出を図った。これは、「交通事故」と「地区交通」から重複する「交通」を取り出すような作業であり、最小のキーワードを取り出すことがねらいである。第3段階では、「土木工学ハンドブック」シソーラス(図-3)に使用されている専門用語を辞書として使用しキーワードの抽出を行った。以上のプロセスで、キーワード群を抽出したが、前述の通り、かな漢字分離では取り出せないキーワードを発見するため、再び論文表題からキーワード群の語句を消去し、残りの文字列から担当者の判断で新たなキーワードを抽出した。本研究のフローで唯一人間の判断を必要としたプロセスである。かなを含むキーワード群を表-1に示す。最終的に得られたキーワード群は、4,835件の論文に対し9,677語であった。

3. 研究キーワードとその複合構造

最初に、約1万語のキーワードの使用頻度を分析した。使用頻度の高い上位50語を表-2に示す。表中、×印を付けた用語は一般的な用語と考えられるため以下の分析には使用しない。その結果、最も使用頻度の高いキーワードは「交通」であり、全論文数の約2割に相当する1,087件の論文に使用されている。次に、「モデル」596件、「道路」557件と続いている。ところで、これらの用語はあまりに基本的用語であり抽象的である。そこで、これらの用語が最終的にどのような用語(文字列)として使用されているかを知るために複合的な構造を分析した。

図-4にその結果の一部を示す。この図は、使用論文数5件以上のキーワードに対して作成し、いずれも結果の一部抜粋である。見方は、「交通」を例に取ると、「交通」を含む1,087件の論文の内、その複合語である「交通量」を含む論文が179件、さ

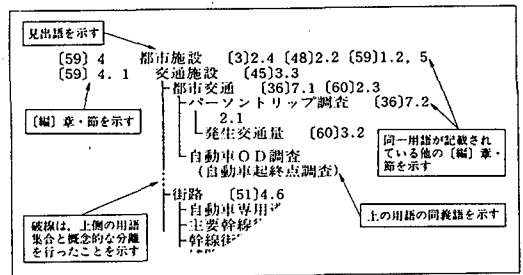


図-3 シソーラスの内容(文献3)より転載)

表-1 かな混じりキーワード群

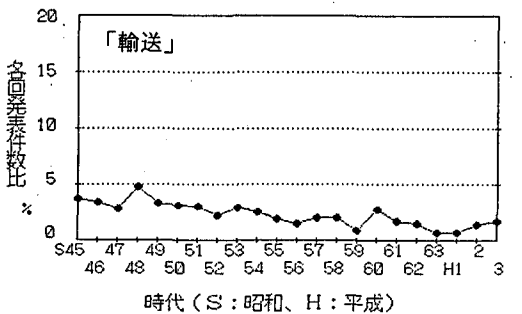
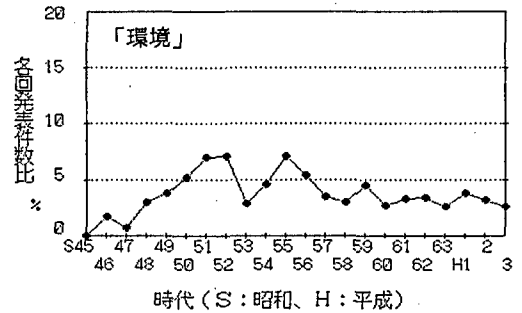
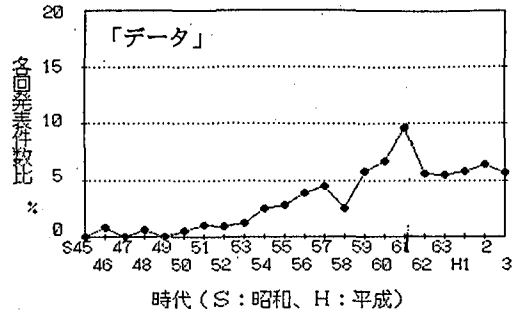
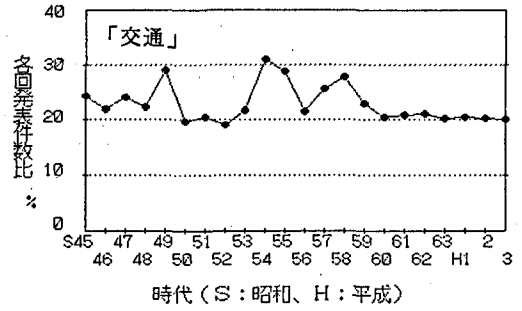
あいまい、あいまいさ、うるささ、きつき、くりかえし、けた、こう束、こう配法、ごみ、すべり、まさつ、ずれ、せせらぎ、たわみ、つき固め、つけ値、でこぼこ、ねじり、のり面、はしけ、バス離れ、ばね、はんげた、ひずみ、ふとう、ふ頭、まくらぎ、ます、まち、まちづくり、ゆらぎ、ゆれ、よこれ、わだち、わだちばね、わだち路、位置づけ、一戸建て、音並み、荷さばき、街づくり、街並み、角折れ、割り振り、割振り、割振り法、巻き込み、巻込み、橋りょう、狂い、狂い進み、狂い値、曲げ、曲上げ、繰返し、計画づくり、結びつき、建て替え、建て物、建べい、建べい率、見積り、見直し、見通し、鋼まくらぎ、山登り法、四国のみち、枝分れ、斜め、車止め、遮へい棟、取り扱い、取扱い、取組み、取締り、手がかかり、手詰まり、手続き、集じん機、住みかえ、住みやすさ、住みよさ、住み替え、重ね合わせ、重み、重みづけ、重み法、出力付き、松枯れ、乗り移り、乗り継ぎ、乗り入れ、乗換え、織り込み、織込み、信号遅れ、水さい、性格づけ、正確さ、積み替え、積合せ、切り取り、切換え、切込み、折れ、折れ角、線ばね、線引き、組み合わせ、組合せ、組合せ、早切り、相乗り、送り型、足掛け、待ち、待ち行列、待ち合せ、待ち合わせ系、待ち時間、値上げ、地すべり、地域づくり、地滑り、遅れ、鋼鉄まくらぎ、追越し、通り、通り狂い、通行止め、低ばね、締め固め、締固め、鉄けた、鉄まくらぎ、鉄まくら木、踏み切り、読取り、突固め、日掃り、入り込み、入込み、買い物、抜け道、飛び出し、表通り、不ばさ、付け値、変わり目、歩み、埋込み式、目づまり、両開き (全153語)

表-2 研究キーワードと使用件数

交通 1,087	環境 182	街路 127
× 研究 1,062	交通量 179	× 推定 124
モデル 596	× 基礎的 177	地区 123
× 考察 588	× 影響 169	発生 123
道路 557	整備 166	高速道路 122
× 分析 552	× 方法 164	× 検討 120
都市 521	× 開発 162	行動 118
計画 421	鉄道 160	バス 117
× 一考察 384	× 解析 152	自動車 115
× 評価 378	選択 151	配分 114
× 利用 334	高速 150	
システム 331	土地 149	
× 特性 299	意識 146	
地域 288	× 基礎的研究 145	
× 予測 241	× 効果 142	
構造 220	施設 141	
× 基礎 198	景観 138	
× 手法 193	分布 130	
× 調査 186	需要 129	
データ 185	時間 128	

・4,835件中、使用頻度の高い上位50語を掲載
・×印は一般的な用語

交通 1087	交通量 179	交通量配分 35	交通量配分モデル 6
			交通量配分法 5
		ト OD交通量 18	
		ト 交通量予測 14	
		ト 分布交通量 13	
		ト 交通量推計 11	
		ト リンク交通量 10	観測リンク交通量 6
		ト 交通量推定 9	
		ト 配分交通量 9	
		ト 発生交通量 8	
		ト 路上交通量 8	
		ト 交通量変動 6	
		ト 自動車交通量 6	
		ト 発生集中交通量 6	
	道路交通 74	道路交通騒音 22	
		ト 道路交通振動 16	
		ト 道路交通流 5	
	交通手段 68	交通手段選択 68	交通手段選択行動 5
		ト 交通手段別 10	交通手段別分担率 9
		ト アクセス交通手段 5	
		ト 通勤交通手段 5	
	交通事故 62	交通事故発生 15	
		ト 交通事故分析 6	
	交通流 61	道路交通流 5	
		ト 交通流モデル 8	
		ト 交通流解析 7	
	交通機関 59	交通機関選択 17	交通機関選択モデル 7
		ト 交通機関別 11	交通機関別分担 9
		ト 交通機関分担 10	交通機関分担モデル 5
		ト 公共交通機関 7	



モデル 596	予測モデル 44	需要予測モデル 13	交通需要予測モデル 6
	選択モデル 41	機関選択モデル 12	交通機関選択モデル 7
		ト 手段選択モデル 9	
		ト 経路選択モデル 8	
	集計モデル 24	非集計モデル 23	
	計画モデル 23		
	行動モデル 23	非集計行動モデル 8	
	配分モデル 21	交通量配分モデル 6	
	モデル化 19		
	分布モデル 18		
	ロジットモデル 16	集計ロジットモデル 7	
	シミュレーションモデル 15		
	評価モデル 15		
	分担モデル 15	機関分担モデル 9	交通機関分担モデル 5
	推計モデル 14		
	土地利用モデル 13		
	立地モデル 12		
	分析モデル 11		
	モデル分析 10		

道路 557	高速道路 122	都市高速道路 39
		└ 高速道路利用 6
		└ 高速道路トンネル 5
		└ 高速道路合流部 5
		└ 首都高速道路 5
	道路網 94	道路網計画 11
		└ 道路網容量 10
		└ 幹線道路網 9
		└ 道路網交通 9
		└ 道路網交通需要推計 8
	道路交通 74	道路交通騒音 22
		└ 道路交通振動 16
		└ 道路交通流 5
		└ 幹線道路網 9
	道路整備 19	
	道路ネットワーク 16	
	道路計画 14	
	道路トンネル 14	高速道路トンネル 5
	地区道路 11	

図-4 キーワードの複合構造

図-5 主要キーワードの使用率の推移

らにその複合語である「交通量配分」を含む論文が35件、そして「交通量配分モデル」が6件あることを現している。「モデル」においては「予測モデル」に関する研究が最も多く、「道路」においては「高速道路」に関する研究が最も多い。さらに具体的なキーワードを知りたければ、より下位のレベルを調べれば良い。こうして、無秩序であった約1万語のキーワードは系統的、構造的に整理していくことが可能となった。ただし、現時点では文字列の照合を行っているだけであり、各キーワードの意味の関連性に踏み込んだ分析には新たな展開が必要である。

4. 研究キーワードの使用頻度の変遷

本章での関心は、前章で得られた主要な研究キーワードの使用頻度を時系列的に分析することにある。ただし、発表件数は回を追う毎に増えているので、論文数そのものの推移を比較しても特徴は現れにくい。そこで、各回の発表件数に対する比率を比較することにする。

図-5は、主な研究キーワードである「交通」、「データ」、「環境」、「輸送」を使用している論文の比率の推移を示したものである。「交通」は、一時的な増減はあるにせよ全体としては常に20%以上の論文に使用されている。「データ」は明らかに増加傾向にある。「環境」は昭和50年代前半に多く使用されその後は停滞気味である。「輸送」は、比率そのものが低いが、傾向としては減少傾向にある。

使用頻度の高い研究キーワード上位100語を対象にその使用傾向を分析した結果、①増加傾向にあるキーワード、②一時的に使用が増えたキーワード、③使用率が比較的安定しているキーワード、④使用率が減少傾向にあるキーワードのグループに分けることができた。結果を表-3にまとめる。1970年から1991年までの22年間においては、「データ」、「モデル」、「システム」、「ファジィ」など情報系の用語が次第に増加しているのが特徴的である。また、環境問題を背景とした「環境」や研究者が高い関心を示した「非集計」などが一時的に多く使用されている。一方、意味としては「環境」にも含まれると考えられるが、「騒音」などは次第に関心が薄れているようにも思われる。

表-3 主要キーワードの使用傾向

傾向	主なキーワード				
増加傾向	・データ ・鉄道 ・ネットワーク	・モデル ・選択	・システム ・時間 ・ファジィ	・整備 ・行動	
一時的増加傾向	・環境 ・災害	・集計 ・ランドサット	・航空	・非集計	
安定傾向	・交通 ・構造 ・施設 ・街路 ・配分 ・自動車	・道路 ・高速 ・景観 ・地区 ・通勤 ・高速道路	・都市 ・土地 ・分布 ・発生 ・軌道 ・土地利用	・地域 ・意識 ・需要 ・バス ・交通量	
減少傾向	・輸送	・計画	・騒音		

6. まとめ

以上、本研究では、論文表題からキーワードを抽出し分析することにより、計画学研究の主要な分野構成や研究者の時代別の関心を明らかにしてきた。分析に使用した年講論文は、1991年（平成3年）から学術情報センターデータベースGAKKAIに登録されるようになり、データベース構築の個人的負担は解消した。今後は、より一層、データベースの有効利用に関心が向けられるべきであろう。

本研究は、著者らが取り組んでいる土木計画における文字データの情報化に関する研究の一環であり、最終的には文献情報の有効利用および土木計画関連用語の体系化を目指したものであることを付記する。

<参考文献>

- 1) 中岡良司・森 弘：OCRシステムによる印刷情報の電算化について、土木学会北海道支部論文報告集、第50号、1994.2
- 2) 中岡良司・森 弘・五十嵐日出夫・佐藤馨一：文献データベースによるキーワード分析の適用、第16回土木情報システムシンポジウム論文集、1991.10
- 3) 土木学会：「第4版 土木工学ハンドブック」、技法堂、1989.11
- 4) クラウス・クリッペンドルフ：「メッセージ分析の技法－内容分析への招待－」、頸草書房、1989.8